



Pattern recognition methods for querying and browsing technical documentation

Karl Tombre, Bart Lamiroy

► To cite this version:

Karl Tombre, Bart Lamiroy. Pattern recognition methods for querying and browsing technical documentation. 13th Iberoamerican Congress on Pattern Recognition, Sep 2008, Havana, Cuba. pp.504-518. inria-00321537

HAL Id: inria-00321537

<https://hal.inria.fr/inria-00321537>

Submitted on 15 Sep 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pattern recognition methods for querying and browsing technical documentation

Karl Tombre¹ and Bart Lamiroy²

¹ LORIA – INRIA, 615 rue du jardin botanique, 54600 Villers-lès-Nancy, France
Karl.Tombre@inria.fr

² LORIA – Nancy Université, B.P. 239, 54506 Vandœuvre-lès-Nancy CEDEX, France
Bart.Lamiroy@loria.fr

Abstract. Graphics recognition deals with the specific pattern recognition problems found in graphics-rich documents, typical technical documentation of all kinds. In this paper, we propose a short journey through 20 years of involvement and contributions within this scientific community, and explore more precisely a few interesting issues found when the problem is to browse, query and navigate in a large and complex set of technical documents.

Key words: graphics recognition, symbol recognition, document analysis, information spotting

1 Introduction

Document image analysis was one of the very first applications of pattern recognition and even of computing. But until the 1980s, research in this field was mainly dealing with text-based documents, including OCR (Optical Character Recognition) [1], and page layout analysis, which is a necessary component for people to be able to build complete page readers [2, 3]. A few people were looking at more specific documents such as music sheet, bank cheques or forms.

At that time, a small community started looking at line drawings [4]. Although the original data were usually scanned documents, the problems were not exactly those of text documents: the aim, as often stated at that time, was to *recognise the high-level objects represented by the graphics*: rebuild the high-level design from an engineering drawing, recognise cities, roads, rivers and regions on a map, input a scanned electrical diagram into an electrical CAD system, etc. This subfield of document image analysis has been known under the name of *Graphics Recognition* since the early 1992s; it is devoted to graphics-rich documents and the specific pattern recognition problems raised by these documents (raster-to-graphics conversion, text-graphics separation, symbol recognition...) and has its own specialized workshop, the GREC series of workshops.

We have been working within this area for many years. Ten years ago, the first author presented an invited paper in Lisbon summarising his results and the remaining open problems after his first ten years in the area [5]. After having rounded the cape of 20 years in the area, it is again time to step back, look at

the achievements, analyse the state of the art, and draw some conclusions for the next years.

This is what the present paper attempts to do, by inviting the reader to a journey through the authors' main contributions to graphics recognition. We then review various interesting problems occurring when dealing with the general issue of being able to browse, index and retrieve information from a large set of graphics-rich documents. As we will see, key points are to be able to take into account the context and to learn in an incremental way.

2 A Short Journey through 20 Years of Contributions to Graphics Recognition

As we had read our classical textbooks on image processing and pattern recognition, we knew that an image analysis process usually involves two main steps:

1. retrieve features from the image through some kind of *segmentation* process;
2. apply some structural analysis and/or some statistical classification to the extracted features to provide *recognition* capabilities.

2.1 Segmentation

In a purely graphical world, it quickly came to our mind that a basic feature was the vector, so as to go from a pixel-by-pixel view of the image to viewing the data as an assembly of geometric entities: line segments, and sometimes circular arcs. This process of retrieving vectors from thick lines in an image, called *vectorization*, may seem a trivial problem, as the skeleton essentially captures the main elongated shapes, which can then be approximated by polygons. But it quickly appears that artefacts, irregularities, noise, junctions, etc. need more elaborate algorithms to avoid having a user correcting “by hand” all the small vectorization errors. Our first published algorithms tried to deal with this by taking a broader view of the picture so as to “look past” the small noise: we matched the polygonalized contours of the shape [6], or we split the image into a grid of small buckets in which we tried to detect simple vectorial configurations [7].

These methods still had a number of parameters to fine-tune for specific applications. There was also a need for being able to detect both vectors and circular arcs. We therefore decided to work out more robust methods, with as few parameters as possible and stable settings [8]. We reconsidered the whole vectorization process [9], using whenever possible state of the art, off-the-shelf methods, and ended up designing a completely new method, based on robust statistical estimators working on the skeleton and aiming at maximum robustness with minimum knowledge [10].

In many cases, the input data are not only graphics but mixed documents containing text, thin lines, thick lines, etc. Therefore, it is not enough to retrieve vectors; we also have to work on the separation of the different layers of raw information. Various methods have been proposed for this task. Following the

general idea of choosing off-the-shelf, stable methods, we started by using the bottom-up approach proposed by Fletcher & Kasturi [11], which analyzes the distribution of connected components in a document on a few simple and generic criteria. But the method has a few limitations, in the case of text connected to the graphics and in its inability to distinguish a dash from a “T”, for instance, and we proposed some improvements [12], without solving the general problem, which is to some extent a “chicken-and-egg” question.

2.2 Recognition

Basing ourselves on the segmentation tools described previously, and on a structural pattern recognition approach, we built complete recognition systems in various areas:

- Cadastral maps with the REDRAW system [13]. In this early work, contextual knowledge about the city maps we wanted to analyze was embedded in the recognition process, through simple heuristics on the kinds of features we were interested in: hatched areas representing buildings, closed areas with labels corresponding to land parcels, streets, etc. Simple, ad hoc procedures recognized hatching and tried to provide higher-level interpretations, with limited success and no attempts at validating the approach to large-scale problems.
- Engineering drawings with the CELESSTIN system [14], where domain knowledge in mechanical engineering was described using a blackboard system. By describing contextual knowledge up to the functional level, we obtained very interesting results on a small subset of drawings representing gears, but the method is barely scalable to a larger area of engineering, due to the complexity of the knowledge representation framework and of the interpretation mechanisms.
- Architectural drawings where we designed a hierarchical symbol recognition system [15] based on the propagation of geometrical and topological constraints in a network and a complete system going all the way to building a simplified 3D model of the represented house [16]. Once again, the method achieved very interesting results but is hardly scalable, both in terms of robustness to noise and in terms of computational complexity.

2.3 Deadlock: Segmentation vs. Recognition

Actually, stepping back to look at these various achievements, of which we were very proud when we reached them, we must admit that although they contributed to progress in the know-how about graphics recognition methods, they did not provide the breakthrough necessary to deal with real-life problems having large numbers of input data with a reasonable variability and complexity.

One of the problems which has already been mentioned is that of the methods’ scalability. Methods can be good at recognizing one graphical symbol among

50, but what about the task of recognizing one symbol among 500 or 1000, especially in the presence of noise?

Another problem is the correction and editing cost. If the recognition rate of a method is 98%, the next question is how important it is to correct manually the 2% which were not recognized. First, you have to find the errors, then you must correct them, if this is deemed necessary. In many cases, this editing must be done and is costly enough for an automatic recognition method not to be economically interesting; the choice will rather be to provide semi-intelligent tools to help trained users, possibly in low-salary countries, input the correct data “by hand”.

Still another problem is the choice of features in the recognition process. Vectors, arcs and hypothetical text blocks are often too low-level to be used directly in recognition processes. In this spirit, our team has worked on various signatures [17, 18] which can capture more information to be used by classifiers or other recognition methods.

But our feeling is that the most important problem is that segmentation itself does not always provide complete enough information for the recognition task to be completed. And in order to segment better, we would need some contextual information provided by recognition. Thus, we stumble upon that which has sometimes been named the Sayre paradox, inspired by Kenneth Sayre’s early work on handwriting recognition [19]: in order to correctly recognize you need to segment, but in order to segment you need to recognize!

Actually, this leads us to the question of how the *context* is taken into account. All recognition tasks do not belong to the same category with respect to the presence of contextual knowledge. In the following sections, we will give three examples of how various types of contexts can lead to various solutions helping us in breaking the segmentation vs. recognition deadlock.

3 Example 1: Closed Context and Hierarchical Recognition

One of the situations where we encountered such a deadlock situation came from a real-world industrial problem, namely that of faultless identification of a set of known symbols in aircraft electrical wiring diagrams³. In this project, funded by the European Commission and involving several companies and research institutions, our main task was to design a recognition method able to distinguish between roughly 500 different symbols. As shown in Fig. 1, some of these symbols can be very similar in shape and only differ by slight details, such as symbols 198 and 199, but they can also be completely different from a visual point of view. Symbols may also be composed of other known and significant symbols and need not necessary be connected (e.g. symbol 208).

The application context of the project required the analysis of very large images of which a huge amount had to be processed. The principal difficulty is

³ The work described in this section was done in collaboration with Jean-Pierre Salmon and Laurent Wendling.

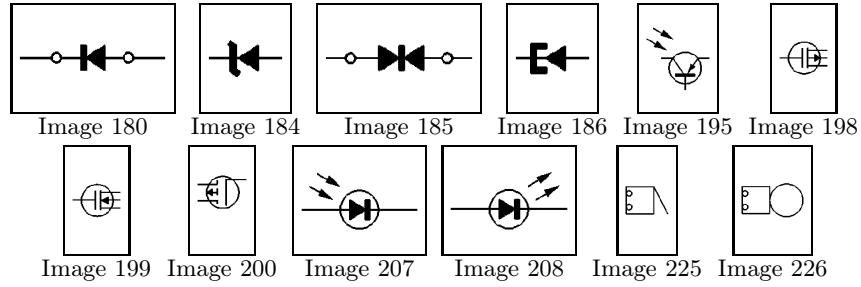


Fig. 1. Examples of images from an electric wiring component database – courtesy Algo'tech Informatique as part of FRESH project funded by European Commission.

to remain scalable with respect to the data volumetry: structural relationships need to be sufficiently expressed in order to proceed to semantic interpretation and eventual electric simulations from wiring diagrams; the approach needs to be sufficiently robustness with respect to image perturbations, but capture detail as well; details should not hinder scalability, etc.

3.1 Segmentation to Localize Symbols

The first issue consists in identifying zones where potential symbols may be located (see also section 4). As wiring diagrams comply with a number of standards and rules, notably the fact that connecting wires are generally horizontal and vertical lines, simple contextual rules were used to implement a semi-automated selection process able to isolate potential symbols.

The remaining problem is now to correctly and rapidly identify the symbols within a set of 500 possible candidates.

3.2 Learning Techniques and Context

State-of-the art learning techniques and global signal-based shape descriptors are usually not well suited to accommodate with composed symbols, nor are they adapted for capturing small detail changes. On the positive side, this tends to make them quite fault tolerant to image distortions. They scale very well and allow for easy adjunction of new symbols in the database.

One could argue that a two step approach might be suitable in this context: first eliminate most of the visual candidate classes, using global signal-based descriptors, and, second, refining the recognition on the identified category of symbols.

This is only valid when preliminary segmentation has been successful in pruning spurious visual distractors, when defining visual classes of similar symbols makes sense, and if there are sufficient data examples on which a classifier (or multiple classifiers) can be trained. But it becomes difficult to achieve when we

have only scarce training data for some of the classes. Furthermore, and more importantly, this approach to learning and classification is quite rigid if the interpretation context changes.

This is specially true in the present case, where the following constraints are met: symbol representative candidates are often unique, and may be subject to considerable visual variations. This makes the cost of statistical learning techniques too expensive, since, in order to create a significant testbed, ground-truth generation (i.e. manual segmentation and identification – see § 5) would mean to solve the problem by hand, given the lack of a sufficient large set of appropriate data. At best, one might consider a user feedback loop to accelerate the manual identification process.

On the other side, structural based shape descriptors are very segmentation dependent and tend not to scale easily when they are based on graph representations, or lack robustness of expressive power when hierarchically organised. However, they capture details very well and provide interesting hooks to semantic interpretation.

The solution we have adopted is to describe symbols as a vector of occurrences of medium level of visual characteristics. In a certain sense, we have created a visual vocabulary, and symbols are considered as an enumeration using this vocabulary.

3.3 Constructing a Visual Vocabulary

The composite descriptor resulting from our work is robust, compact and fast, and composed of smaller elementary descriptors. These elementary descriptors (the vocabulary) are thoroughly evaluated with respect to robustness and reliability and have integer or boolean values. They express information like the number of small occlusions, the occurrence of text, the number of connected components, the existence of symmetrical extensions, presence of full circles, rectangles, etc. The composite merged descriptor obtained in this way is compact, fast, discriminant, scalable and error tolerant.

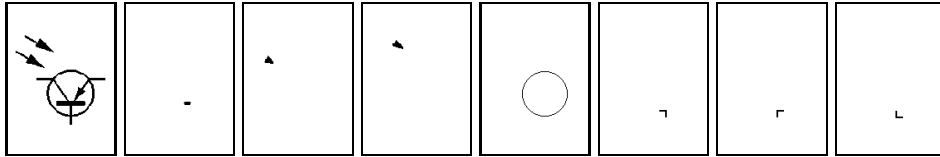


Fig. 2. Extraction of Robust Structural Vocabulary: initial symbol, filled forms and circles, oriented corners

One of the solutions is to express visual information using a vocabulary constructed on direct extraction of (more or less) complex structures in the images. This vocabulary is based on currently unpublished work (inside project reports of the FRESH project) that characterizes symbols by a set of very robust, local

structures. These structures need not necessarily be extracted by a structural or syntactic methods. In our case, for instance, our group has developed specific and specialised detectors for circles [10, 20], oriented corners, loose endpoints, rectangles, etc., as shown in Fig. 2.

These descriptors end up describing a symbol as a vector of their occurrences, and symbols are considered as being similar when their respective descriptor vectors are similar. On a testset of approximately 250 symbols, Fig. 3 gives an indication of how discriminant our method is, using 13 descriptors.

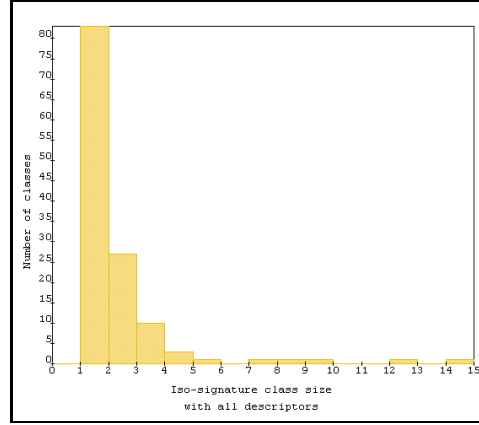


Fig. 3. Iso-distance class histogram for 234 symbol classes, and 13 descriptors

3.4 Partial Conclusion on the Use of Context

The point here is not to evaluate the approach with respect to others, but to consider how context has been used to achieve the classification and recognition goals. We are here in the presence of a perfect illustration of how closed-form context information and classification/recognition is considered in most main stream state-of-the art approaches. Closed-form approaches consider that the set of visual classes is finite and well defined (in our case, for instance, we have a known set of 500 symbols) and their appearance or visual context is constrained (e.g. in our case, connected in wiring diagrams, no rotation, limited scale factors, etc.). As a matter of fact, the context of use is implicitly embedded in the approaches themselves and is hardly formalized. The global approach follows in most cases the following line:

- gather a sufficiently exhaustive set of context-coherent ground truth examples;

- either apply learning techniques, unsupervised or supervised classification, or human expert knowledge to conceive a *decision ruleset*⁴ for classifying, selecting, recognition, etc.
- assess the quality of the obtained *decision ruleset* over another set of data that implicitly respects the same context.

This general approach has the advantage of being robust and generic, in the sense that it can be applied as is to a wide variety of configurations and contexts. It is however very rigid where the context itself is concerned, and requires that it can be captured (implicitly or explicitly) by a sufficiently exhaustive example set that has been previously established.

4 Example 2: Open Context, Bootstrapping without Full Recognition: Symbol Spotting

A second case in which we can find an issue to the segmentation vs. recognition paradox is that of information spotting. In the previous example, the context was closed and known and the purpose was to extract all the available information from the document. But there are a number of applications where the need is *not* for full-scale recognition of the documents, but rather for localization of some useful information without any claim at being able to analyze the whole content.

This is sometimes called *information spotting* and has been addressed in various application contexts in document image analysis. One typical example is that of historical documents. Due to their complexity and to the difficulty of being able to build a complete analysis and recognition process, several authors have proposed word or text spotting mechanisms for this kind of documents [21–24]. Bertrand Couiasnon et al. essentially used the same idea when they built a platform for providing access by content to handwritten documents such as old military registers [25]. Other examples include detecting low contrast strings in complex tables [26], locating the title and author regions in document images [27], or robust detection of text in very noisy document images [28].

In graphics recognition, several researchers have also started looking at ways to get information from maintenance drawings [29], extract useful information from the title block of technical drawings [30], or detect symbols in maps [31], without going through a complete recognition process.

Symbol spotting methods rely on various pattern recognition methods working on features such as region-adjacency graphs [32], vectorial signatures [33], or signatures based on force histograms [17] or on the Radon transform [18], which enable us to localize and recognize complex symbols in line-drawings.

When it comes to spotting target symbols, structural approaches are powerful in terms of their representational capabilities. In our group, Zuwała used a simple structural representation of symbols to introduce a hybrid approach

⁴ By lack of any better term, “decision ruleset” describes any algorithm that, given a set of ground-truth data, and any given learning or classification method, is capable of address the needs within the implicit context of use covered by the selected data.

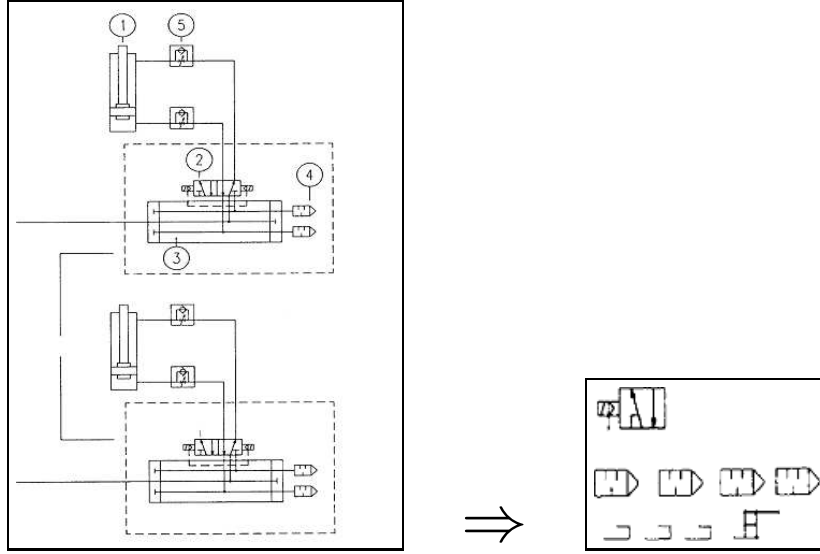


Fig. 4. Example of symbol spotting on an engineering drawing, from [34]. The user delineates a symbol (left) and a number of candidates are retrieved (right).

for processing symbols connected to other graphical information [34, 35]. Fig. 4 illustrates the working of the system: when a candidate symbol is selected in the document, a number of candidate regions are retrieved.

5 Example 3: Learning the Context Provided by the User

In the previous sections we have been considering the fact that pattern recognition methods rely on a more or less implicit presence of context of use. Context information is used to:

1. define the set of patterns (or pattern classes) to recognize,
2. find the most adequate description for achieving their recognition,
3. develop the adapted segmentation or localization techniques to bootstrap recognition, or to define the area on which to compute the descriptors,
4. provide hooks to higher level interpretation processes.

As a matter of fact, this is what *context* is all about: it defines what is visually adequate, how it may be distorted, what may be surrounding it (be it clutter or other significant visual information), and what it refers to in a higher representation level. Up to now, most pattern recognition techniques have been focussing on “*Given a context, how do I recognize?*”. But what about context changes, or missing information?

5.1 Relevance Feedback

One of the most obvious difficulties is the absence of a clearly defined training set defining what to recognize. Since classical recognition techniques, as mentioned previously, are based on the computation of a pattern descriptor and a metric between descriptors, the only need is to have a reference descriptor (or a set of reference descriptors, or a manifold of reference descriptors ...) describing the object classes that are to be recognized. However, due to inevitable tolerance to deformation and form variability, there is a strong need to define how to define the correct metrics to use. This is usually done, either by using invariant descriptors, when deformations are small and well controlled, or by statistic learning techniques, provided sufficient ground truth is available.

Unfortunately, finding a descriptor and a metric that maps an “absolute” concept of similarity into a distance measure between low level features is nearly impossible since we have no means of determining what intra-class deformation is acceptable. Depending on the context of use, the same user, using the same documents, may have different goals, and thus implicitly require different intra-class variability.

One of the solutions is therefore to ask the human user to define, by means of relevance feedback, to iteratively assess the quality of a dynamically computed metric or classification decision rule. Early relevance feedback systems were built using heuristic-based techniques where the main idea was to estimate an *ideal query point* maximizing the correlation between the user similarity concept and the low level image features. Further developments formalized the problem to minimization of the total distance between the positive samples to the query [36, 37]. The principal findings were that the optimal query is obtained by averaging the positive samples, and that the Mahalanobis distance is the optimal weighted metric. MindReader [36] and Mars [38] CBIR systems apply these approaches with success.

Actual state-of-the-art tends to consider the problem as a two class classification issue and tries to adapt known classification schemes to take into account the supplementary difficulties resulting from the small number of training samples, and the asymmetry in the data set. For instance, Su [39] adapts a bayesian classifier based on the maximum likelihood, estimating the boundary of the relevant items from the positive samples and assigning penalties to unlabeled samples close to a negative one. While Zhang [40] and Onada [41] use techniques based on Support Vector Machine, trying to iteratively determine the best hyperplane separating the positive and the negative samples in the projection space.

Other experiments have been carried out, involving decision trees [42, 43] or nearest neighbors [44]. Work in our group [45, 35] has used the latter to construct an unconstrained navigation tool in handwritten documents.

5.2 Concept Learning and Analysis

Another possible (and prospective) direction to develop is to try and use a basic visual vocabulary that is as independent as possible of segmentation, and image

deformations (see Fig. 2 in § 3.3) and to use learning techniques to discover what is visually pertinent and how it is organised. Once the basic vocabulary is defined, we can obtain a straightforward way of describing the image that combines both expressiveness and very high flexibility. On the one hand one can reduce or extend the size of the vocabulary in function of what robust descriptors are available. On the other hand expressing relationships based on their relative positioning [46–49] provides the tools of creating very rich pattern descriptions.

The remaining problem is how to explore what this new representation can offer in terms of recognition, classification or learning of concepts. We are going to explore this in the next section, by feeding these data to a Inductive Logic Programming process.

What is different with this approach is that it introduces a new dimension to the learning that has traditionally been considered in the broad image analysis domain. Most learning, clustering and classification techniques used for recognition purposes can essentially be boiled down to smart co-occurrence and correlation discovery approaches. They essentially fail to include higher order relationships.

Inductive Logic Programming (ILP) combines automatic learning and first order logic programming. By encoding relationships between visual primitives as first order predicates, the system is able to learn quite different notions than what could be done by “regular”, statistical learning techniques. For instance, we have used the vocabulary defined in section 3.3, and used simple predicates to describe relative positioning of visual components. By selecting image set {195, 198, 199, 200, 207, 208} as representatives of one sort of symbol, and the others are counter examples, the ILP solver⁵ gives the following result:

```
symbol(A) :-
    has_element(A,B), type(B,circle), has_element(A,C), inside(C,B),
    type(C,cornernw).
```

which, translated into natural language, means that the chosen examples all have circles containing a north-west corner element (◐). This is something statistical, pixel or primitive based approaches will have a hard time to discover, since the exact (nor even relative) position of the corner is not defined. Better, the system could as well have discovered rules like

```
symbol(A) :-
    has_element(A,B), type(B,circle), has_element(A,C), inside(C,B).
```

where any symbol containing a circle containing some other element would fit, regardless its shape, size or position. Given a shape vocabulary and a set of relationships, the approach is able to assess which combination and which level of precision is appropriate for the context being considered.

⁵ We have been using Aleph, freely available from the Oxford University Computing Lab at <http://web2.comlab.ox.ac.uk/oucl/research/areas/machlearn/Aleph/aleph.html>.

5.3 What Does This Tell us About Context ?

In this last section we have been exploring how parts of the implicit context knowledge could be learnt or otherwise represented; by human feedback on the one side, or by formal learning techniques on the other side, based on a predefined, but extensible vocabulary including relative positioning predicates. This gives us a first order logic based description of the symbols, on which ILP can be used to extract “semantic” contexts or concepts.

As straightforward as this might seem, this is actually a profound paradigm shift. Instead of “*Given a context, how do I develop the best toolbox for recognition ?*” as described through the general state-of-the-art approaches, and specifically the one in section 3, the question here is: “*Given a toolbox, what can I recognize ?*” (replace *toolbox* by *vocabulary, knowledge, context* or whatever suits you best). Using human interaction and relevance feedback, this consists in evaluating how discriminant the tuple description + classification model can be, eventually revealing fundamental limitations that need to be overcome or investigated. The introduction of more formal methods like the one using ILP described in the previous section takes this a step further, by opening up classification and allowing that the description of the symbols be easily mixed with other information. The main advantage of this approach is that this information needs not necessarily be visually represented (for example, it may come from surrounding text or other co-occurrent meta-data or even multimedia content), and that it thus opens a new scope of possible combined text/image concept characterisation and learning.

Further work will probably need to focus on extending the initial vocabulary (e.g. introducing the notion of connectivity, refining the inclusion predicate or using relative distance and size). These are all straightforward extensions that are readily available from an image analysis standpoint. More prospective work will want to connect this to Formal Concept Analysis [50] and Galois Lattices as to automatically cluster the symbol space without having to define, as in this paper, which symbols belong to a particular class, and see if this corresponds to a visual reality.

6 Conclusion

Graphics recognition methods allow us to extract useful information from graphics-rich documents. Whereas the possibility of building analysis system able to deal 100% automatically with a large documentation remain probably at the level of wishful thinking, graphics recognition tools can still be used in very useful applications where the aim is to extract sufficient information to be able to browse the documentation, index it according to some useful features, and spot symbols and other areas of interest.

In this paper, we have given hints at ways to take into account the context, and if possible learn it from user input. After 20 years of activity in this field, a number of achievements have been reached but there are still many exciting challenges for the new generations of researchers.

References

1. Mori, S., Suen, C.Y., Yamamoto, K.: Historical Review of OCR Research and Development. *Proceedings of the IEEE* **80**(7) (July 1992) 1029–1058
2. Tang, Y.Y., Lee, S.W., Suen, C.Y.: Automatic Document Processing: A Survey. *Pattern Recognition* **29**(12) (December 1996) 1931–1952
3. Nagy, G.: Twenty Years of Document Image Analysis in PAMI. *IEEE Transactions on PAMI* **22**(1) (January 2000) 38–62
4. Smith, R.W.: Computer Processing of Line Images: A Survey. *Pattern Recognition* **20**(1) (1987) 7–15
5. Tombre, K.: Ten Years of Research in the Analysis of Graphics Documents: Achievements and Open Problems. In: *Proceedings of 10th Portuguese Conference on Pattern Recognition*, Lisbon, Portugal. (March 1998) 11–17
6. Tombre, K.: Finding and Coding Graphics in a Composite Document. In: *Proceedings of 5th Scandinavian Conference on Image Analysis*, Stockholm (Sweden). (1987) 589–596
7. Vaxivière, P., Tombre, K.: Subsampling: A Structural Approach to Technical Document Vectorization. In Dori, D., Bruckstein, A., eds.: *Shape, Structure and Pattern Recognition (Post-proceedings of IAPR Workshop on Syntactic and Structural Pattern Recognition*, Nahariya, Israel). World Scientific (1995) 323–332
8. Tombre, K., Ah-Soon, C., Dosch, P., Habed, A., Masini, G.: Stable, robust and off-the-shelf methods for graphics recognition. In: *Proceedings of the 14th International Conference on Pattern Recognition*, Brisbane (Australia). (August 1998) 406–408
9. Tombre, K., Ah-Soon, C., Dosch, P., Masini, G., Tabbone, S.: Stable and robust vectorization: How to make the right choices. In Chhabra, A.K., Dori, D., eds.: *Graphics recognition—Recent advances*. Volume 1941 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin (September 2000) 3–18
10. Hilaire, X., Tombre, K.: Robust and Accurate Vectorization of Line Drawings. *IEEE Transactions on PAMI* **28**(6) (June 2006) 890–904
11. Fletcher, L.A., Kasturi, R.: A robust algorithm for text string separation from mixed text/graphics images. *IEEE Transactions on PAMI* **10**(6) (1988) 910–918
12. Tombre, K., Tabbone, S., Péliissier, L., Lamiroy, B., Dosch, P.: Text/graphics separation revisited. In Lopresti, D., Hu, J., Kashi, R., eds.: *Proceedings of the 5th IAPR International Workshop on Document Analysis Systems*, Princeton, NJ (USA). Volume 2423 of *Lecture Notes in Computer Science*, Springer-Verlag (August 2002) 200–211
13. Antoine, D., Collin, S., Tombre, K.: Analysis of Technical Documents: The RE-DRAW System. In Baird, H.S., Bunke, H., Yamamoto, K., eds.: *Structured Document Image Analysis*. Springer-Verlag, Berlin/Heidelberg (1992) 385–402
14. Vaxivière, P., Tombre, K.: Celesstin: CAD Conversion of Mechanical Drawings. *IEEE COMPUTER Magazine* **25**(7) (July 1992) 46–54
15. Ah-Soon, C., Tombre, K.: Architectural Symbol Recognition Using a Network of Constraints. *Pattern Recognition Letters* **22**(2) (February 2001) 231–248
16. Dosch, P., Tombre, K., Ah-Soon, C., Masini, G.: A complete system for analysis of architectural drawings. *International Journal on Document Analysis and Recognition* **3**(2) (December 2000) 102–116
17. Tabbone, S., Wendling, L., Tombre, K.: Matching of Graphical Symbols in Line-Drawing Images Using Angular Signature Information. *International Journal on Document Analysis and Recognition* **6**(2) (2003) 115–125

18. Tabbone, S., Wendling, L., Salmon, J.P.: A new shape descriptor defined on the Radon transform. *Computer Vision and Image Understanding* **102**(1) (April 2006) 42–51
19. Sayre, K.M.: Machine recognition of handwritten word: A project report. *Pattern Recognition* **5**(3) (1973) 213–228
20. Lamiroy, B., Gaucher, O., Fritz, L.: Robust Circle Detection. In: *Proceedings of 9th International Conference on Document Analysis and Recognition, Curitiba (Brazil)*. (2007) 526–530
21. Manmatha, R., Rothfeder, J.L.: A Scale Space Approach for Automatically Segmenting Words from Historical Handwritten Documents. *IEEE Transactions on PAMI* **27**(8) (August 2005) 1212–1225
22. Leydier, Y., Lebourgeois, F., Emptoz, H.: Text search for medieval manuscript images. *Pattern Recognition* **40**(12) (December 2007) 3552–3567
23. Rath, T.M., Manmatha, R.: Word spotting for historical documents. *International Journal on Document Analysis and Recognition* **9**(2–4) (April 2007) 139–152
24. Konidaris, T., Gatos, B., Ntzios, K., Pratikakis, I., Theodoridis, S., Perantonis, S.J.: Keyword-guided word spotting in historical printed documents using synthetic data and user feedback. *International Journal on Document Analysis and Recognition* **9**(2–4) (April 2007) 167–177
25. Coüasnon, B., Camillerapp, J., Leplumey, I.: Access by content to handwritten archive documents: generic document recognition method and platform for annotations. *International Journal on Document Analysis and Recognition* **9**(2–4) (April 2007) 223–242
26. Li, Y., Wang, Z., Zeng, H.: Correlation Filter: An Accurate Approach to Detect and Locate Low Contrast Character Strings in Complex Table Environment. *IEEE Transactions on PAMI* **26**(12) (December 2004) 1639–1644
27. Xiao, Y., Yan, H.: Location of title and author regions in document images based on the Delaunay triangulation. *Image and Vision Computing* **22**(4) (April 2004) 319–329
28. Zheng, Y., Li, H., Doermann, D.: Machine Printed Text and Handwriting Identification in Noisy Document Images. *IEEE Transactions on PAMI* **26**(3) (March 2004) 337–353
29. Syeda-Mahmood, T.: Indexing of Technical Line Drawing Databases. *IEEE Transactions on PAMI* **21**(8) (August 1999) 737–751
30. Najman, L., Gibot, O., Barbey, M.: Automatic Title Block Location in Technical Drawings. In: *Proceedings of 4th IAPR International Workshop on Graphics Recognition, Kingston, Ontario (Canada)*. (September 2001) 19–26
31. Samet, H., Soffer, A.: MARCO: MAP Retrieval by COntent. *IEEE Transactions on PAMI* **18**(8) (August 1996) 783–798
32. Locteau, H., Adam, S., Trupin, E., Labiche, J., Héroux, P.: Symbol Spotting using Full Visibility Graph Representation. Presented at 7th IAPR International Workshop on Graphics Recognition, Curitiba (Brazil) (September 2007)
33. Rusiñol, M., Lladós, J.: Symbol Spotting in Technical Drawings Using Vectorial Signatures. In Liu, W., Lladós, J., eds.: *Graphics Recognition—Ten Years Review and Future Perspectives (Revised Selected Papers from GREC’05)*. Volume 3926 of *Lecture Notes in Computer Science*. Springer-Verlag (2006) 35–46
34. Zuwala, D., Tabbone, S.: A Method for Symbol Spotting in Graphical Documents. In Bunke, H., Spitz, A.L., eds.: *Document Analysis Systems VII: Proceedings of 7th International Workshop on Document Analysis Systems, Nelson (New Zealand)*. Volume 3872 of *Lecture Notes in Computer Science*. (February 2006) 518–528

35. Zuwala, D., Rendek, J.: Browsing graphics without prior knowledge. In: Proceedings of the 18th International Conference on Pattern Recognition, Hong-Kong (China). (2006) 735–738
36. Ishikawa, Y., Subramanya, R., Faloutsos, C.: MindReader: Query databases through multiple examples. In: Very Large Databases. (1998)
37. Rui, Y., Huang, T.: Optimizing Learning in Image Retrieval. In: Computer Vision and Pattern Recognition. (June 2000) 1236
38. Rui, Y., Huang, T., Mehrotra, S.: Content-Based Image Retrieval with Relevance Feedback in MARS. In: Proceedings of IEEE International Conference on Image Processing. (1997) 815–818
39. Su, Z., Zhang, H., Ma, S.: Using Bayesian Classifier in Relevant Feedback of Image Retrieval. In: 12th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'00). (2000)
40. Zhang, L., Lin, F., Zhang, B.: Support vector machine learning for image retrieval. In: Proceedings of IEEE International Conference on Image Processing. (2001) 721–724
41. Onada, T., Murata, M., Yamada, S.: Relevance feedback document retrieval using support vector machines. In: Proceedings of International Joint Conference on Neural Networks (IJCNN-2003). (2003) 1757–1762
42. MacArthur, S.D., Brodley, C.E., Shyu, C.: Relevance Feedback Decision Trees in Content-Based Image Retrieval. In: IEEE Workshop on Content-based Access of Image and Video Libraries. (2000) 68
43. Wang, T., Rui, Y., Hu, S., Sun, J.: Adaptive Tree Similarity for Image Retrieval. *Multimedia Systems* **9** (2003) 131–143
44. Giacinto, G., Roli, F.: Bayesian relevance feedback for content-based image retrieval. *Pattern Recognition* **37**(7) (2004) 1499–1508
45. Rendek, J., Lamiroy, B., Tombre, K.: A Few Step Towards On-the-Fly Symbol Recognition with Relevance Feedback. In Bunke, H., Spitz, A.L., eds.: Document Analysis Systems VII: Proceedings of 7th International Workshop on Document Analysis Systems, Nelson (New Zealand). Volume 3872 of Lecture Notes in Computer Science. (February 2006) 604–615
46. Matsakis, P., Wendling, L.: A New Way to Represent the Relative Position Between Areal Objects. *IEEE Transactions on PAMI* **21**(7) (1999) 634–643
47. Matsakis, P., Keller, J.M., Sjahputera, O., Marjamaa, J.: The Use of Force Histograms for Affine-Invariant Relative Position Description. *IEEE Transactions on PAMI* **26**(1) (January 2004) 1–18
48. Bloch, I.: Fuzzy Spatial Relationships for Image Processing and Interpretation: a Review. *Image and Vision Computing* (23) (2005) 99–110
49. Bennett, B., Agarwal, P.: Semantic categories underlying the meaning of ‘place’. In: Spatial Information Theory: Proceedings of the 8th International Conference (COSIT-07). Volume 4746 of Lecture Notes in Computer Science., Springer-Verlag (2007)
50. Ganter, B., Stumme, G., Wille, R., eds.: Formal Concept Analysis: Foundations and Applications. Volume 3626 of Lecture Notes in Computer Science. Springer-Verlag (2005)